



Research papers

Combining signal decomposition and deep learning model to predict noisy runoff coefficient

Arash Rahi^{a,c,*}, Mehdi Rahmati^{b,c}, Jacopo Dari^{a,d}, Heye Bogena^c, Harry Vereecken^c, Renato Morbidelli^a

^a Dept. of Civil and Environmental Engineering, University of Perugia, via G. Duranti 93, 06125 Perugia, Italy

^b Department of Soil Science and Engineering, Faculty of Agriculture, University of Maragheh, Maragheh, Iran

^c Institute of Bio and Geosciences (IBG), Forschungszentrum Jülich, 52428 Jülich, Germany

^d National Research Council, Research Institute for Geo-Hydrological Protection, via Madonna Alta 126, 06128 Perugia, Italy

ARTICLE INFO

Handling Editor: C. Corradini

Keywords:

Long Short-Term Memory
Discrete wavelet transform
Runoff coefficient
Soil water storage
Total evaporation

ABSTRACT

Predicting runoff coefficient (Rc), as an indicator of the catchment's response to the rainfall-runoff process, remains a persistent challenge using different modelling techniques, especially in catchments with strong human manipulation. This study investigates the efficiency of the Long Short-Term Memory (LSTM) method in predicting Rc for the Rur catchment, in Germany. The period from 1961 to 2021 is considered, which is subject to human intervention and significant urbanization especially in the northern part of the catchment. An LSTM structure is defined by employing inputs at a monthly resolution including temperature, precipitation, soil water storage, and total evaporation with a look-back window of 1 to 6 months to model noisy Rc data of the study area. Two approaches using either undecomposed or decomposed Rc were employed in conjunction with the LSTM method, to mitigate the impact of noise associated with Rc. The results show that in the case of undecomposed Rc, the best performance of the LSTM structure was obtained with a 4-month look-back window, yielding Nash-Sutcliffe efficiency (NSE) of 0.55, 0.46, and 0.15 for training, validation, and test sets, respectively. These results highlight inadequate accuracy in accounting for the presence of noise in Rc. Therefore, in the second novel approach, we used maximal overlap discrete wavelet transform (MODWT) to decompose the Rc up to level 3 to reduce the complexity and distribute the noise effects across each level. The new approach showed high accuracy in modelling noisy data of Rc with NSE values of 0.97, 0.95, and 0.90 for training, validation, and test sets, respectively. The obtained results underscore the pivotal role of decomposition techniques in conjunction with LSTM to account for the presence of noise, especially in catchments with strong human manipulation.

1. Introduction

The runoff coefficient (Rc) was initially formulated by Sherman (1932) in the early twentieth century. It represents the ratio of runoff to rainfall (Bedient et al., 2008; Burak et al., 2021) and is extensively employed in engineering practices (Chen et al., 2020). In hydrology, Rc serves as a diagnostic variable, playing a key role in describing runoff generation within a catchment (Merz et al., 2006), and which is well-known to be influenced at different scales by the ongoing global warming (Blöschl et al., 2019; Masseroni et al., 2021; Dari et al., 2023). Understanding the impact of climatic and human-induced changes on local hydrological processes is essential, emphasizing the significance of

assessing Rc changes in river basins (Velpuri and Senay, 2013). The Rc displays substantial variability in both space and time due to various influencing factors (Norbiato et al., 2009; Merz and Blöschl, 2009; Penna et al., 2011). For example, Chen et al. (2020) highlighted the crucial influence of groundwater levels on the temporal variability of the Rc. Merz and Blöschl (2009) studied the variation of Rc across a broad range of Austrian basins and found that Rc are mainly influenced by climate indicators like mean annual precipitation and the long-term ratio of actual evaporation to precipitation, with minimal impact from land use, soil types, and geology in the studied catchments. Similarly, Massari et al. (2023) investigated how basin pre-storm conditions influence the Rc across 284 European basins and 60,620 flood events.

* Corresponding author at: Dept. of Civil and Environmental Engineering, University of Perugia, via G. Duranti 93, 06125 Perugia, Italy.

E-mail address: arash.rahi@dottorandi.unipg.it (A. Rahi).

<https://doi.org/10.1016/j.jhydrol.2024.131815>

Received 8 March 2024; Received in revised form 20 June 2024; Accepted 31 July 2024

Available online 13 August 2024

0022-1694/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Strong correlations of Rc were observed with deep soil storages, pre-storm discharge, and snow water equivalent among different basins. Their findings emphasized accurate flood forecasting relies on understanding how pre-flood conditions relate to the Rc, to establish links for early-warning systems. Rahi et al. (2023) used continuous wavelet coherence analysis (WCA) to investigate the impact of hydroclimatic variables, including temperature (T), precipitation (P), soil water storage (SWS), and Land Use Land Cover (LULC) changes, on the Rc. Their findings revealed that precipitation correlation with Rc remained moderately unstable. In contrast, weak correlations were observed with LULC changes, while a strong positive but lagged correlation with SWS (1 month) and a strong lagged (3–6 months) but negative correlation with T suggested the prevailing significance of hydroclimatic factors over LULC changes.

The evolving challenges in complex hydrological issues like real-time flood and drought forecasting, as well as water resource management, underscore the need for precise modelling of the regional-scale rainfall-runoff process. Despite numerous proposed approaches and the development of various models over the past decades, a comprehensive solution remains hard to attain. This challenge is driven by the substantial temporal and spatial variability inherent in the rainfall-runoff system (Merz and Blöschl, 2009) and the limitations in the availability of suitable mathematical tools to fully capture the dynamics underlying this process. Artificial Intelligence (AI) techniques are widely used in hydrology due to their flexibility, transferability, and computational efficiency (Wu et al., 2009; Nourani et al., 2011; Shoaib et al., 2014). However, the limitation of more classical AI-based techniques such as Artificial neural networks (ANNs) lies in the absence of temporal information on model inputs, affecting their ability to capture the temporal reaction of runoff to meteorological data (Kratzert et al., 2018). Hydrological hysteresis, a time-dependent phenomenon (Gharari and Razavi, 2018), complicates rainfall-runoff simulations by retaining the historical impacts on present or future occurrences. With the evolution of computational hardware capabilities and algorithms during the past decades, particularly long short-term memory (LSTM) as a class of recurrent neural networks (RNN) has become prominent. The LSTM outperforms the traditional models in simulating runoff, specifically in addressing time series data challenges in hydrology (Kratzert et al., 2018; Bergen et al., 2019; Mao et al., 2021). The LSTM model has exhibited considerable potential in addressing temporal issues by memorizing sequence, accounting for prolonged temporal relationships and multiple output variables (Xiang et al., 2020). LSTM's key advantage lies in its capacity to capture and learn long-term dependencies between the input and output data within the network (Kratzert et al., 2018). Systems that show the presence of hysteresis or of which the evolution depends on past states are typically denoted as “path-dependent systems” or “systems with memory” (Hassani et al., 2014). However, hydrological hysteresis has not been extensively studied in the context of water circulation (Gharari and Razavi, 2018), some studies indicated that there might be a lagged response of runoff to meteorological variables, which can extend beyond 100 days (Chen and Kumar, 2002; Orth and Seneviratne, 2013). In this regard, a range of studies has explored the use of LSTM models in capturing temporal relationships in wind power and flow stream estimation using static look-back windows (Wang et al., 2020; Giriagama et al., 2022), while Li et al. (2021) introduced dynamic time-delays to improve the prediction performance of industrial time series models.

Standard RNNs face long-term dependence issues, such as gradient vanishing or exploding (Hochreiter and Schmidhuber, 1997). Despite efforts with variants like LSTM, gated recurrent unit (GRU), bidirectional LSTM (BiLSTM), and bidirectional GRU (BiGRU), these models may struggle with highly nonlinear and noisy time series without proper preprocessing (Zhou et al., 2024). Nhita et al. (2016) proposed to use the moving average as an effective data smoothing method, enhancing the modelling and forecasting capabilities when using neural networks. Another approach to attenuate noise involves decomposing the problem

into smaller subproblems, reducing complexity, and distributing the noise effects across each subproblem for more effective mitigation (Sáez et al., 2014). Recent advances in hybrid machine learning models (MLMs) have utilized various time series decomposition techniques to improve the MLM performance, including discrete wavelet transform (DWT) (Mallat, 1989; Liu et al., 2014), maximal overlap DWT (MODWT) (Adamowski and Sun, 2010), wavelet packet transform (WPT) (Moosavi et al., 2017), empirical mode decomposition (EMD) (Huang et al., 1996; Napolitano et al., 2011), and ensemble EMD (EEMD) (Wang et al., 2015) and variational mode decomposition (VMD) (Dragomiretskiy and Zosso, 2014). These methods decompose complex time series data into simpler sub-series for separate modelling, effectively capturing valuable information at different temporal scales (Gokhale and Khanduja, 2010). In addition, Adamowski and Sun (2010) showed that combining an ANN with DWT improved the prediction of daily streamflow. Furthermore, Seo et al. (2017) proposed VMD-based extreme learning machine (VMD-ELM) and VMD-based least squares support vector regression (VMD-LSSVR) models, which demonstrated higher accuracy for daily rainfall-runoff modelling compared to single and DWT-based MLMs. Finally, Guo et al. (2023) found that hybrid models combining two-stage decomposition methods with LSTM, specifically WT with CEEMDAN, VMD, and LMD, outperformed traditional and standalone models in forecasting daily runoff over the Pearl River in China.

The MODWT decomposes a signal in both time and frequency domains, similar to the DWT. However, unlike the DWT, the MODWT retains down-sampled values at each decomposition level and can be applied to signals of any length without the power-of-two restriction. According to Cornish et al. (2006), the MODWT offers several advantages over the DWT. These advantages suggest that combining MODWT with soft computing models could be a more effective and efficient approach for river stage modelling (Seo et al., 2017).

Despite extensive studies on Rc and its influencing factors, there remains a gap in modelling Rc at regional and global scales. This paper addresses this gap by utilizing LSTM to predict Rc, incorporating the decomposition method. The first configuration of undecomposed Rc has been considered to investigate the capability of the LSTM method in modelling noisy Rc. The second configuration is defined to mitigate noise effects in Rc modelling, using the MODWT method to decompose Rc into sub-components, thereby reducing noise impact and improving model accuracy.

2. Study area

The analysis has been carried out using data from the Stah outlet of the Rur catchments, in Germany (see Fig. 1). It covers an area of 2,245 km² predominantly situated in the North Rhine-Westphalia region and is part of the TERENO (TERrestrial ENvironmental Observatories) network in Germany (Zacharias et al., 2024). Small portions of the basin extend into Belgium (6.7 %, 157 km²) and the Netherlands (4.6 %, 108 km²) as well. The terrain elevation gradually decreases from south to north, ranging from 680 to 30 m above sea level. The annual precipitation ranges from 650 to 1200 mm/yr increasing from north to south of the catchment (Bogena et al., 2005a). Fig. 1b provides land cover features at 100 m resolution derived by the Corine Land Cover data set referring to the year 2018 (CLC 2018). Agricultural areas cover 49 % of the catchment, mainly concentrated in the lowland regions of the northern part, while urban areas cover approximately 17.7 % of the area, primarily concentrated in the same northern region. Natural areas are predominantly situated in the elevated regions of the catchment, accounting for 31.8 % of the entire Rur catchment area. Water bodies and peat bogs represent less than 1.5 % of the total area (Fig. 1b). The upland region (southern part), characterized by low permeable rock strata and limited groundwater recharge, relies on reservoirs for managing water resources. The upper Rur catchment exhibits pronounced seasonality, displaying a fast runoff response following rainfall events (Rudi et al., 2010). In contrast, water management strongly regulates runoff in the

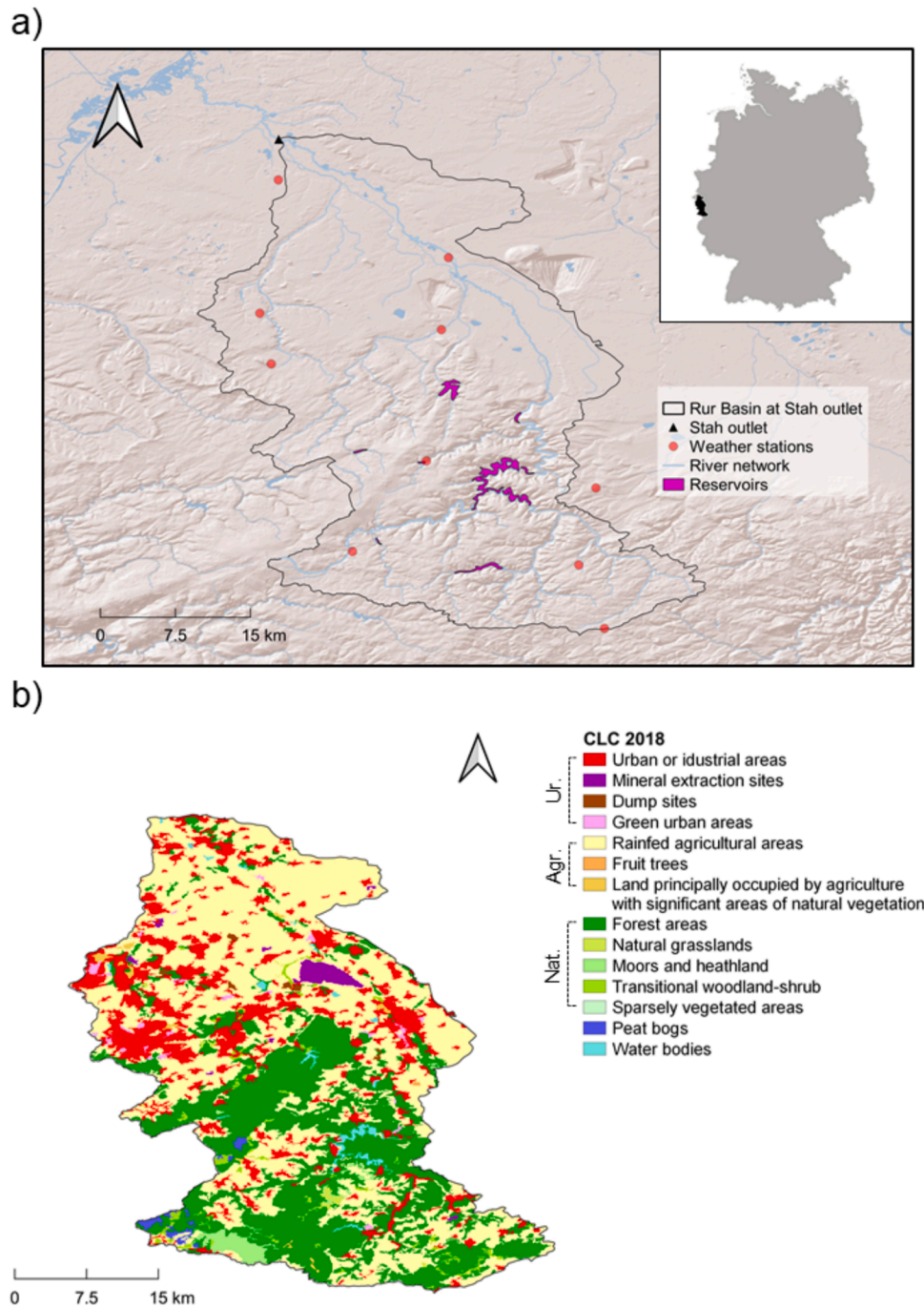


Fig. 1. Study area map. Panel a) provides an overview of the Rur catchment (the ESRI Shaded Relief layer is used as a background map), the monitoring network, and the basin location with respect to Germany, and panel b) shows land cover features derived from the Corine Land Cover data set referring to the year 2018 (CLC 2018). Ur., Agr., and Nat. refer to urban, agricultural, and natural areas, respectively.

lower Rur catchment (Bogena et al., 2005a). The flatland region from the centre to the north, containing loose rock aquifers, features relatively high groundwater recharge rates (Bogena et al., 2005b), which are subject to intensive groundwater withdrawal. The Rur is a typical low mountain river with a very unbalanced discharge behaviour (the ratio of discharge values can fluctuate between 1:1000 and 1:2000 in extreme cases of dry weather and flood discharge). To compensate for these strong fluctuations, as well as for drinking water and energy production, a networked water reservoir system with several dams was established in the first part of the 20th century in the upper reaches, with a maximum reservoir volume of around 300 million cubic meters.

3. Materials and methods

3.1. Data

In this investigation, monthly runoff volumes recorded at the Stah outlet during the period 1961–2021 were used. The runoff height was subsequently computed by dividing the volume by the basin's area. For the timeframe 1961–2021, historical P data were downloaded by the climatic data center service (CDC, https://opendata.dwd.de/climate_environment/CDC/), and information was gathered from 10 meteorological stations within the study area (see Fig. 1a). Cumulative monthly values of P recorded by each rain gauge were collected for all considered

years, and the Thiessen polygons method was applied to derive basin-averaged monthly P data. Then, Rc for the entire catchment was calculated as the ratio between runoff height and basin-averaged P.

The average monthly T data at 2 m elevation, the cumulative monthly data of total evaporation (ET), and the average monthly data of volumetric soil water content (SWC) from 1961 to 2021 were acquired from the Copernicus ERA5-Land (European Reanalysis v5-Land) service (Muñoz-Sabater et al., 2021) with a spatial sampling of 9 km. To quantify SWS, the initial 4 layers of the product's physical schematization were considered, encompassing a depth of up to 2.89 m, SWS represents the total amount of water that is stored in the soil profile. All the data were spatially averaged over the Rur basin prior to the analyses (Table 1).

Overall, the choice to use in-situ rainfall data ensures that we leverage the high-resolution, long-term observational data available in the region, thus enhancing the reliability of our LSTM forecasts. This approach aligns with our study's focus on combining LSTM modelling and wavelet decomposition to improve forecast accuracy in an anthropized basin covering 2,245 km².

The primary focus of this study was to develop a technique to overcome the effects of noise while forecasting Rc. The proposed technique can be potentially implemented elsewhere, upon availability of adequate records to fit the train set. For larger scale Rc forecasts, combining AI-based model with wavelet technique using satellite data can even be applied, though this is beyond the scope of our current research.

3.2. LSTM modelling

LSTM, a specialized form of RNN, overcomes the limitations of traditional ANNs in managing long-term dependencies through the integration of memory structures (Hochreiter and Schmidhuber, 1997). In this study, an LSTM structure has been defined to develop a hydrological model for predicting Rc.

3.2.1. LSTM design

The basic model includes a cell state C for each time step, preserving crucial information for short- and long-term memory. Each cell receives the previous state of the target variable h_{t-1} and one or more predictor variables x as input and returns the current state of the target variable h_t as the output. Fig. 2a illustrates the core structure and algorithms of a modern LSTM model, updating six parameters in each time step through equations (1) to (6):

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{c}_t \quad (4)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \times \tanh(C_t) \quad (6)$$

The first parameter, f_t , is associated with the forget gate, determining the extent to which the previous cell state should be disregarded. This decision is made by a sigmoid function with a linear calculation involving the current input x_t and the previous state h_{t-1} . The weights (W) and biases (b) in each LSTM cell vary in different steps for the linear equations. A lower f_t value (close to zero) prompts the sigmoid function to forget the previous cell state C_{t-1} more. The second parameter, i_t , is linked to the input gate, deciding what new information is to be retained and added to the cell state. Calculated by a sigmoid function with a linear relation on x_t and h_{t-1} , it influences the candidate values \tilde{c}_t for the new cell state, determined by a hyperbolic tangent (tanh) function. Subsequently, the cell state C_t is updated. Finally, the output parameter o_t is computed by a sigmoid function with a linear relation on x_t and h_{t-1} . The output result at the current state, h_t , is the product of o_t and the tanh function value of the cell state C_t . In summary, f_t , i_t , \tilde{c}_t , C_t , o_t , and h_t are all vectors. Specifically, f_t and i_t range between 0 and 1, \tilde{c}_t ranges between -1 and 1, C_t can take any real value, o_t ranges between 0 and 1, and h_t ranges between -1 and 1. All these vectors allow the LSTM network to handle multiple units of data simultaneously, which enhances its ability to learn and model complex temporal patterns.

In this study, we set our predictor variable x to contain states of T, P, SWS, and ET as Chen and Xu (2021) demonstrate that the inclusion of more meteorological variables can improve LSTM performance. To frame our model, we employed the pre-existing LSTM layer components within the Keras (Chollet, 2021) framework in Python. In principle, each cell of the LSTM model only takes into account the previous state of the target variable, but not the previous states of the predictors. Therefore, any possible lagged relationship between the predictors and the outcome variable must be considered. In this regard, a new predictor variable is usually obtained from the original predictor variable, using a specific look-back window. In our analysis, we examined different look-back windows with values ranging from 1 to 6 months to obtain the best look-back window according to the trial-and-error concept.

Shukla et al., (2023) compared the runoff simulations and observed data for the calibration period (2000–2010) and the validation period (2011–2015) at the stations Monschau, Linnich, and Stah in the Rur catchment. They found the model performed best at Monschau, likely due to fewer anthropogenic interventions compared to the more urbanized and agricultural northern parts of the basin, where water use, channel diversions, and dams, which are not considered in the model, significantly affect stream flow.

Considering the regulated runoff due to dam construction and water treatment in the northern part of the Rur catchment, the stochastic nature of rainfall data may lead to high variance and noisy Rc signals, complicating modelling efforts. Consequently, the Rc estimated in this work is influenced by those above-mentioned human activities. However, studies (e.g., Shukla et al., 2023) indicate that the Rur river's discharge dynamics remain relatively natural, as less than 5 % of runoff is used for drinking water and efforts are made to maintain natural discharge conditions to protect the ecosystem. Therefore, the Rc values from this study should still approximate natural conditions, though precise quantification of the dam-regulating effect is beyond this study's scope. To overcome the noise effects, we implemented two configurations: 1) we modelled the noisy signals of Rc without any preprocessing and 2) we first decomposed Rc signals using MODWT into low-frequency (A) and high-frequency (D) components (see Table 2) and then modelled each component separately and combined them at output level to reconstruct the original signal of Rc (see Fig. 2c):

$$Rc = \left(\sum_{i=1}^m D_i \right) + A_m \quad (7)$$

where m is the level of decomposition.

Table 1
Data sets used in this study and their main features.

Data	Temporal resolution	Spatial sampling	Data source
P	Monthly, 1961–2021	Point-wise	CDC
Runoff	Monthly, 1961–2021	Point-wise	TERENO-NRW
SWC	Monthly, 1961–2021	9 km	ERA5-Land services
T	Monthly, 1961–2021	9 km	ERA5-Land services
ET	Monthly, 1961–2021	9 km	ERA5-Land services

CDC (Climate data centre): https://opendata.dwd.de/climate_environment/CDC/.

TERENO-NRW (Nord Rhein Westfalen): <https://www.tereno.net/>.

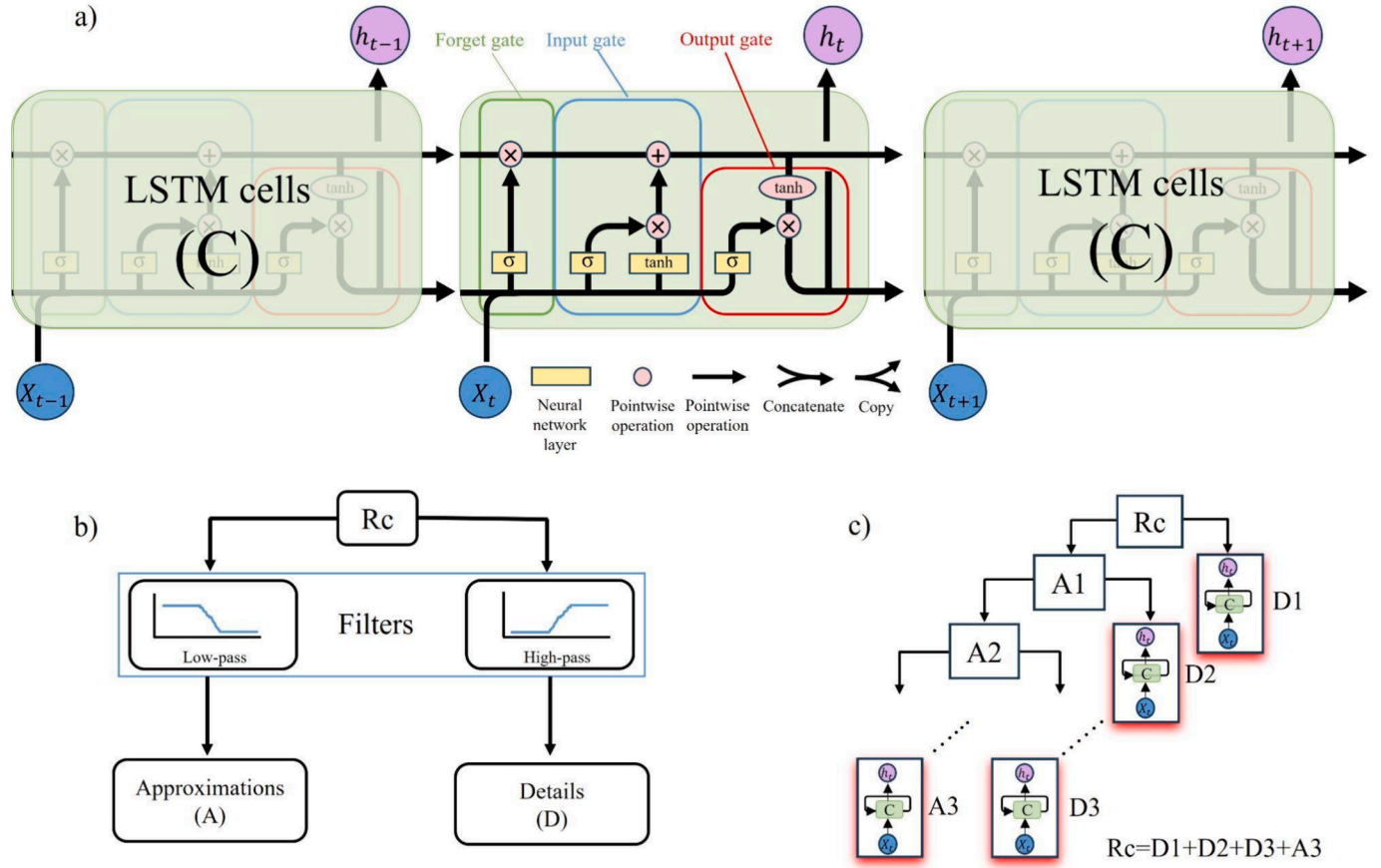


Fig. 2. LSTM structure: a) basic layer configuration from time step $t-1$ to $t+1$, with detailed calculations illustrated in the LSTM cell at time step t , b) schematic of decomposing Rc using MODWT c) LSTM modelling through Rc's wavelet decomposition tree.

Table 2

Decomposition components and corresponding frequencies.

Components	Frequencies (cycles/month)	Relative energy (%)
D1	0.25–0.5	6.62
D2	0.113–0.277	3.2
D3	0.0561–0.136	4.07
A3	0–0.061	86.11

3.2.2. Model setting and parametrization

In the model architecture, two LSTM layers are employed: the first layer consists of 512 neurons with a rectified linear unit (ReLU) activation function and is configured to return sequences. The subsequent LSTM layer contains 256 neurons, also utilizing the ReLU activation function. Following the LSTM layers, a dense layer, with an activation function of ReLU, is added with a number of neurons equal to the target output's shape. An epoch represents one complete pass through the entire training dataset, allowing the model to update its parameters to improve performance. To prevent overfitting, an early stopping indicator is implemented with a patience of 20 epochs, restoring the best weights based on monitoring the validation loss. The optimization process utilizes stochastic objective functions, based on adaptive estimates (Adam) developed by Kingma and Ba (2015), a gradient descent-based algorithm, with an initial learning rate set to 0.0001, and it is adjusted by a factor of 0.001 when the loss on the validation set fails to decrease, thereby mitigating overfitting.

3.2.3. Evaluation metrics

The model aims to minimise the fraction of unexplained variance, employing mean squared error (MSE) as the loss function. The choice of

MSE aligns with the Nash-Sutcliffe efficiency (NSE), a widely-used performance evaluation method in hydrological modelling. NSE ranges from $-\infty$ to 1, and the closer its value is to 1, the better the model performs (Arnold et al., 2012):

$$MSE = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n} \quad (8)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (9)$$

Where Y_i is the observation at the time i , \hat{Y}_i is the model result at time i , \bar{Y} is the mean of all observations, and n is the total number of observations. In alignment with Moriasi et al. (2015), this study utilizes NSE to categorize results into unsatisfactory ($NSE \leq 0.5$), satisfactory ($0.5 < NSE \leq 0.7$), good ($0.7 < NSE \leq 0.8$), and very good ($NSE > 0.8$), providing a concise evaluation of model accuracy and reliability.

Root mean square error (RMSE) serves as an indicator of the standard deviation of the prediction error in a model. A smaller RMSE value indicates a smaller prediction error, signifying better performance of the model (Naabil et al., 2017). RMSE represents the errors between predicted values and actual values and is calculated as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (10)$$

In the current study to verify the hydrological modelling of Rc, two metrics, NSE and RMSE, were used to evaluate the differences between Rc simulated in the train set (1961–2001), validation set (2001–2011) and test set (2011–2021) with observed Rc.

3.2.4. Model optimization

The optimization of the LSTM model parameters was carried out differently for the main parameters and hyperparameters. The look-back window parameter, was tuned by testing and iteration to achieve the optimum output, while the hyperparameters are predefined settings that control the learning process (see Table 3) were optimised by Grid Search Cross-Validation (GridSearchCV), a Hyper-Parameters Optimization (HPO) technique used in machine learning (Goel et al., 2020; Bischl et al., 2023). Grid search is a systematic approach to parameter tuning in machine learning models, evaluating each combination of algorithm parameters specified in a grid. In this process, the key terms include the estimator, implementing the scikit-learn estimator interface, and representing the classifier to be trained. The parameter grid is a Python dictionary with parameter names and corresponding settings, testing all combinations to identify the optimal accuracy. Cross-validation is crucial in this context, determining the resampling strategy for evaluating models and mitigating bias. It involves randomly shuffling and splitting the dataset into k groups, using each as a test set while the others act as training data. This process provides more reliable results compared to a simple train-test split, as it ensures that each sample is present in the testing data once and used for training $k-1$ times (Ranjan et al., 2019).

In this study, several key hyperparameters (see Table 3) have been optimized to achieve the best possible performance of our model. The remaining hyperparameters, which have been set to default or predetermined values based on empirical best practices, include the learning rate and early stopping (see section 3.2.2). This approach ensures a balanced between computational efficiency and model accuracy, allowing us to focus on the most impactful hyperparameters for our specific application.

3.3. Model configurations

As previously explained, two different configurations have been defined to implement and compare the results of the proposed LSTM modelling approach: 1) Undecomposed Rc and 2) Decomposed Rc.

3.3.1. Undecomposed Rc

In this approach, the undecomposed Rc has been modelled by the LSTM method considering a look-back window, which was set on the basis of the concept of “lag” derived by the WCA, a theory recently applied to hydrological studies (Rahmati et al., 2020; Rahi et al., 2023; Dari et al., 2024). Fan et al. (2020), found that an improper window size can significantly impact model performance. Referring to Appendix A, we assessed the phase shift and lag between the benchmark signal (Rc) and P, SWS, T, and ET time series following Rahmati et al. (2020), then different look-back windows evaluated by LSTM to find the optimum value.

3.3.2. Decomposed Rc

In the second approach, the MODWT was used to decompose Rc using the Coif mother wavelet into its approximation (A) and details (D), depending on the decomposition level. The MODWT analyses signals

into progressively finer octave bands. This multiresolution analysis enables to detect patterns that are not visible in the raw data. The level for MODWT computation (decomposition level) is indicated as a positive integer, not exceeding the floor of the $\log_2 N$, where N represents the size of the input layer in the time dimension. In this study, with a time dimension of 720 months (starts from January 1961 to December 2020), the maximum level is set to 9. However, Srivastava et al. (2016) argue that the level of decomposition can be recognized by visual inspection where the approximation coefficient and detail have visual similarity. Therefore in accordance to Srivastava et al. (2016) plus inspection of the accuracy of the developed models, we decided to decompose Rc up to 3 levels. The MODWT partitions the energy of the signal across the various scales and scaling coefficients. The relative energy is defined as the proportion of energy in each component by level, relative to the total energy of decomposition level. The sum of relative energies in all components equals 1 (see Table 2).

Then, the decomposed signals (one approximation and three details) were modelled separately applying the LSTM method with the same look-back window setting as defined in previous section. Finally, the modelled decomposed signals were combined with each other to reconstruct the Rc signal. By comparing the reconstructed and undecomposed Rc signals, we were able to evaluate the accuracy of the model at the output level, in addition to checking the accuracy of each individual model for each component.

4. Results and discussion

The outcomes of this study reveal commendable stability in the monthly resolution performance of the LSTM model, showcasing its ability to reproduce the complex features of the Rc. Based on the WCA results (Appendix A), a temporal lag of 1–6 months between Rc and the key input variables (i.e., P, SWS, T, and ET) has been identified. The temporal lag deduced by the WCA has been considered as a look-back window. After several trials considering entities of the look-back window within the range of 1–6 months, the optimal LSTM model performances have been found when employing a look-back window equal to 4 months. Such a value is a good trade-off between capturing relevant temporal dependencies and preventing overfitting (Fig. 4).

In this study, GridSearchCV was employed to optimize hyperparameters (Listed in Table 3). The findings revealed that the most accurate results in forecasting both decomposed Rc and undecomposed Rc can be obtained by applying the optimized parameters as presented in Table 4. In the following subsections, we will report the results obtained for each configuration.

4.1. Undecomposed Rc

In Fig. 3a results of training and validation loss (look-back window equal to 3 months) show that the model fails to predict undecomposed Rc accurately, and the result contains uncertainty and needs to be improved. The NSE for both the train and validation sets is equal to 0.51 and 0.49 respectively, indicating unsatisfactory match with observed Rc from 1961 to 2001 (Fig. 3b). However, a noteworthy challenge arises in the test set, where the NSE drops to 0.05, indicating diminished accuracy and underscoring the imperative need for a targeted solution. Holman-

Table 3
Input hyper-parameters entered into the GridSearchCV method.

Hyper-parameters	Value
batch size	4, 8, 16, 32, 64
number of layers	2, 3, 4, 5
LSTM units	256, 512, 1024
Epoch	50, 100, 200
optimizers	‘sGD’, ‘Adam’, ‘RMSprop’
activations	‘tanh’, ‘relu’, ‘sigmoid’

Batch size = the number of instances in each training batch.

Stochastic gradient descent (SGD).

Root mean squared propagation (RMSprop).

Table 4
Optimal values of hyperparameters – returned by the GridSearchCV method.

Hyper-parameters	value
batch size	4
number of layers	2
LSTM units	512
Epoch	200
optimizer	‘Adam’
activations	‘relu’

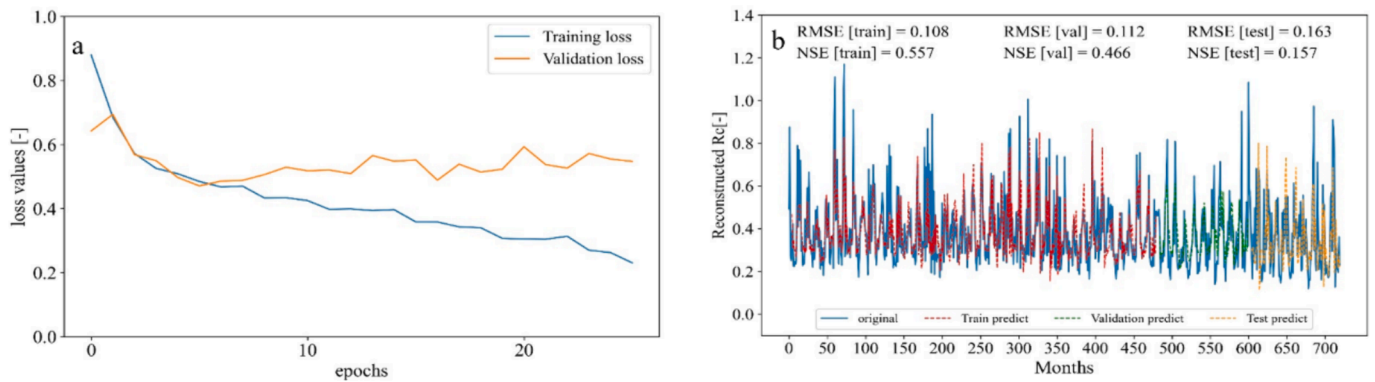


Fig. 3. Results of fitting R_c by LSTM model using look-back window equal to 4 months a) Loss value, b) RMSE and NSE. Months since January 1961.

Dodds et al., (1999) found that the temporal sampling of precipitation can affect model calibration, where lower sampling frequencies may result in the underestimation of runoff. This is probably due to missing the possible lagged responses between precipitation and runoff, which probably is the case in our analysis, too. Therefore, we modelled the undecomposed R_c by the LSTM method applying a look-back window

equal to 4 months. The results showed that the NSE of the train and validation set remained constant, while NSE of test set enhanced to the value of 0.15 (Fig. 3d). Although using the look-back window equal to 4 improved the results, the results were still very poor. Patry et al. (1989) and Gusev et al. (2017) both highlight the role of uncertainty in hydrological models, particularly in the presence of noise. Onstad and

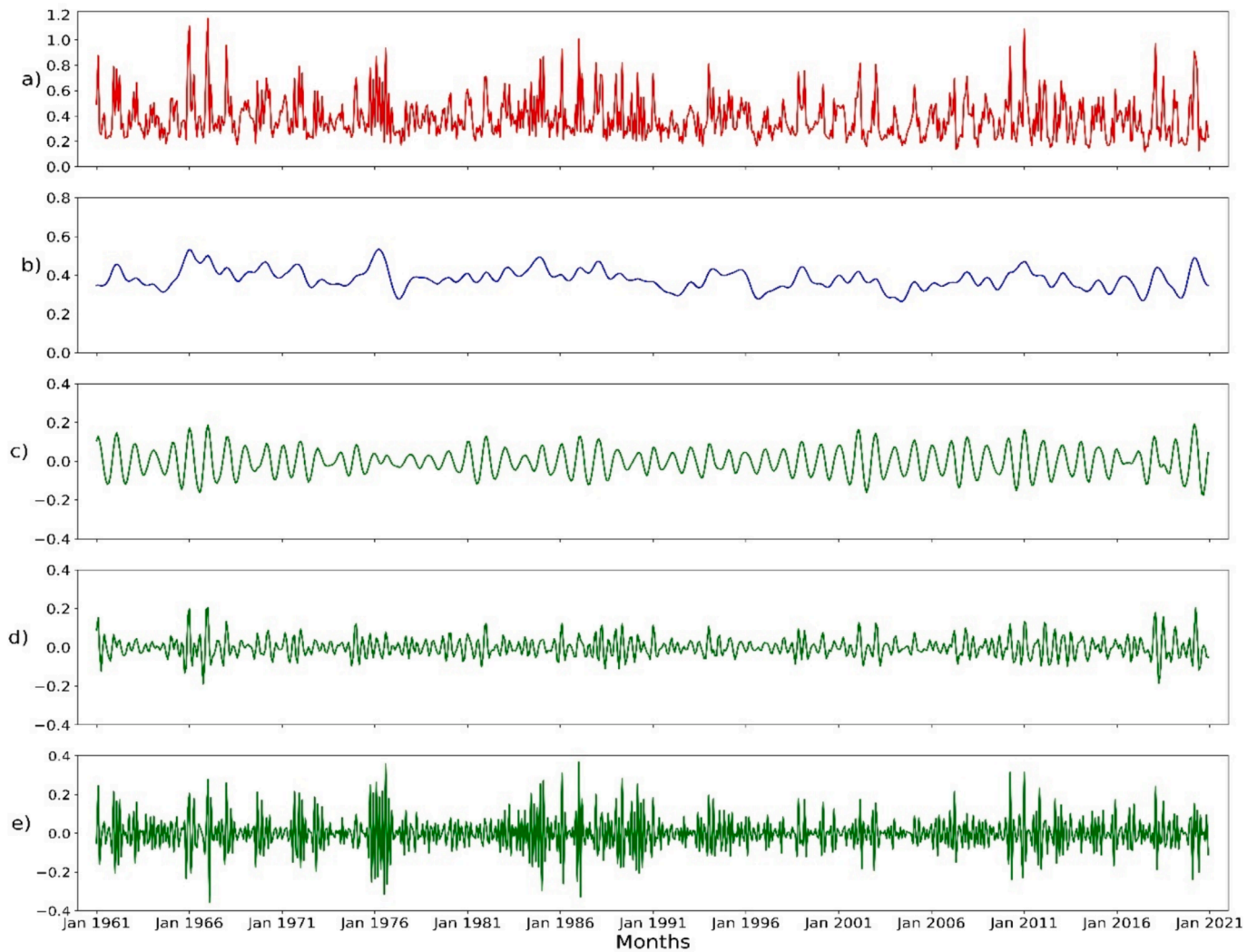


Fig. 4. Decomposed R_c at level 3 by MODWT (Coif1 mother wavelet); a) Original R_c , b) Approximation at level 3 (A3), c) Detail at level 3 (D3), d) Detail at level 2 (D2), e) Details at level 1 (D1); Red, blue and green mark plots refer to original R_c , approximation and details plots respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Jamieson (1970) further emphasize the importance of considering land use modifications in runoff modelling, which can be affected by noise. Gelfan et al. (2015) extend this discussion to the analysis of hydrological extreme uncertainty, which is also influenced by noise. Therefore, upon a detailed examination of the R_c time series, the primary source of discrepancy was traced back to inherent noise associated with precipitation and runoff data. These noise elements introduced a level of unpredictability in the R_c values, leading to suboptimal model fitting during the training phase and subsequently reduced accuracy in forecasting.

4.2. Decomposed R_c

The obtained results for this approach show that the detail component D1 contain varying amounts of noise (panels (e) of Fig. 4), whereas the noise (low frequency) from components D2, D3 and A3 is extremely minimal and indistinguishable (see panels b-d of Fig. 4).

In panel (a) of Fig. 5, the training and validation loss profiles for the first level of decomposed R_c are depicted, which correspond to the highest frequency (0.25–0.5) component in the signal. The presence of high-frequency noises in runoff and precipitation can lead to overfitting in hydrological models. The overfitting of training data compromises the model's suitability to be generalized, rendering its performance unreliable when applied to new measurements (Piotrowski and Napiorkowski, 2013). The result of this study indicates that the training set has been effectively simulated; however, the validation set exhibits signs of overfitting which is attributable to the presence of high-frequency noises in the D1. Examining panel (c) of Fig. 5 reveals NSE values for the train, validation, and test sets equal to 0.92, 0.81, and 0.71, respectively. Furthermore, RMSE values for the train, validation, and test sets are observed as 0.03, 0.03, and 0.05, respectively. Despite the expected presence of high-frequency noises in the D1, the LSTM model demonstrates a well-fitted behaviour, yielding significant simulation results across train, validation, and test data sets. Upon inspecting Fig. 5b, it becomes apparent that the fitting for component D2 surpasses the D1.

Notwithstanding this improvement, the overfitting of the validation set persists, albeit with noticeable advancements. The LSTM model's enhanced fitting component D2 is reflected in the improved NSE values for the train, validation, and test sets, showing as 0.97, 0.80, and 0.81, respectively. This enhancement in the efficacy metrics underscores the model's proficiency in capturing underlying patterns within the data. Concurrently, the RMSE values for the train, validation, and test sets are equal to 0.01, 0.02, and 0.03, respectively. Hence, the model has effectively mitigated the impact of high-frequency noises associated with D1 and D2, showcasing acceptable performance in simulating R_c dynamics across the train, validation, and test datasets.

In Fig. 6a, a notable improvement in modelling accuracy is evident for D3. The NSE values, reaching 0.99, 0.99, and 0.96 for the train, validation, and test sets respectively, underscore the robust performance of the model. Such high NSE values indicate the model's ability to accurately capture the complexity of the data across train, validation, and test datasets. Furthermore, RMSE values equal to 0.005, 0.008, and 0.001 for the train, validation, and test sets, respectively, are found (refer to Fig. 6c).

The modelling of the component A3 using LSTM (refer to Fig. 6b and d) has been implemented with high precision. The training and validation loss presented in Fig. 6b reflects an outstanding simulation, resulting in NSE values of 0.99, 0.99, and 0.99 for the train, validation, and test sets, respectively. This model was expected to exhibit the most precise fitting among component D3 and A3. This anticipation arises from the fact that the approximation coefficient filters by a low-pass frequency, aligning with the characteristics of the mother tongue. The RMSE for the A3 consistently maintains a low value, reminding the behaviour of the D3, which exhibits the least uncertainty among all decomposed details. This stability underscores the robustness of the A3 in capturing underlying patterns with minimal errors, assuring its crucial role in achieving accurate and stable results within the broader context of the decomposition analysis.

Inspecting the efficacy values shows that the LSTM could fit the train set, however, the efficacy does not show satisfactory accuracy value

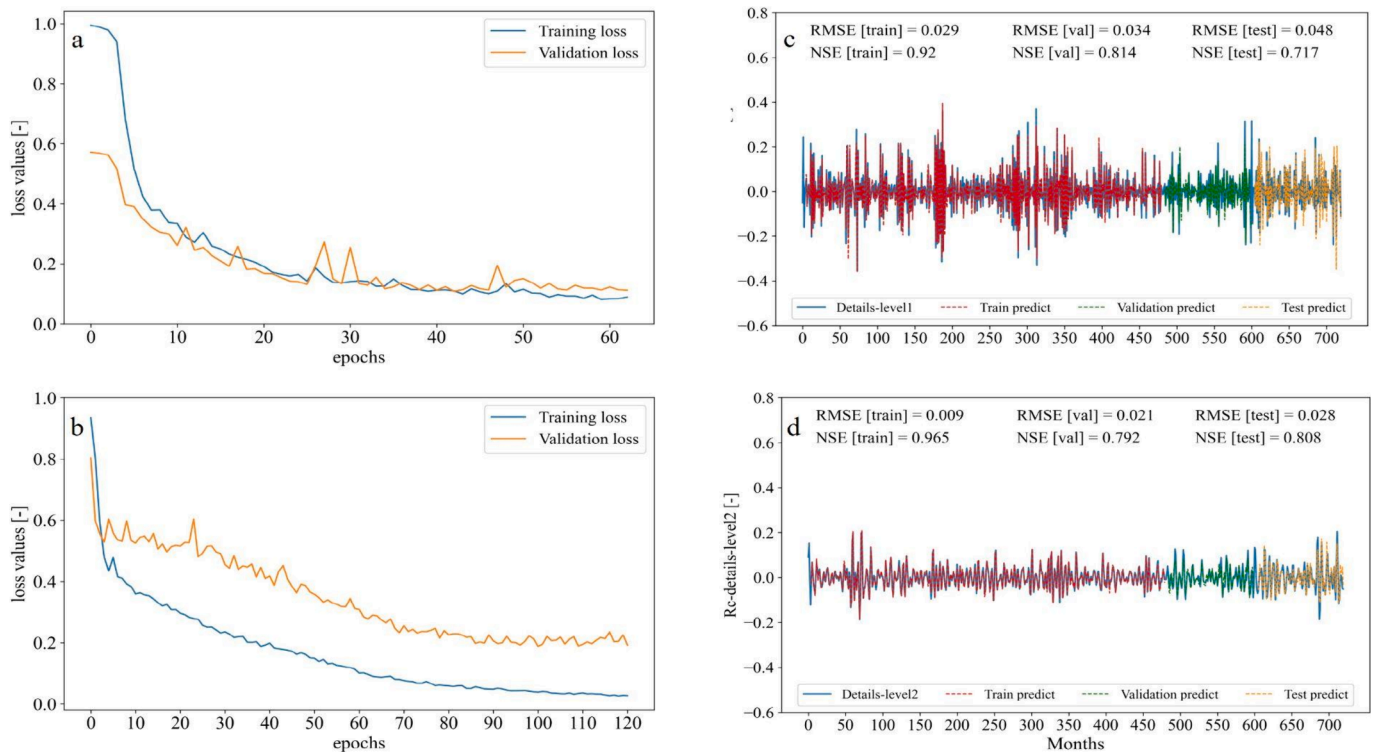


Fig. 5. Results of fitting R_c details by LSTM model (A look-back window = 4) a) Loss value of D1, b) Loss value of D2, c) RMSE and NSE of D1, d) RMSE and NSE of D2. Months since January 1961.

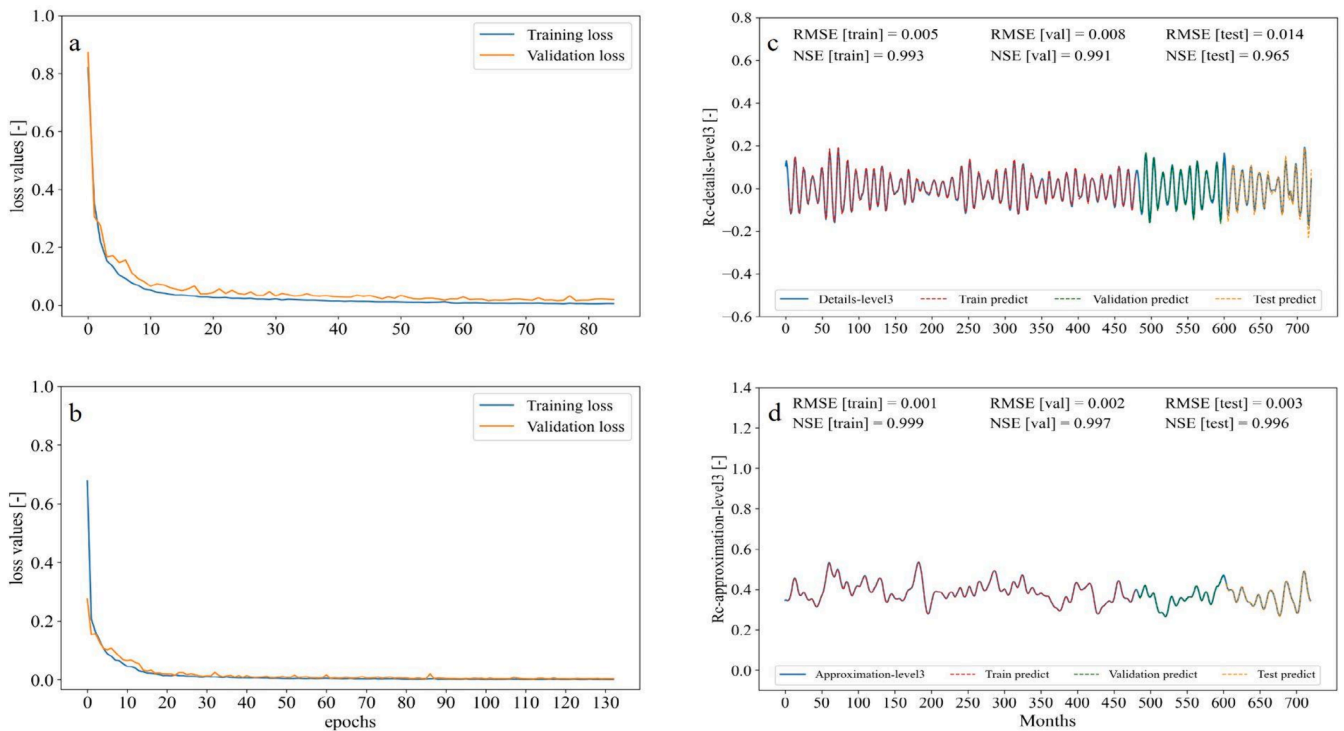


Fig. 6. Results of fitting Rc details by LSTM model (A look-back window = 4) a) Loss value of D3, b) Loss value of A3, c) RMSE and NSE of D3 d) RMSE and NSE of A3. Months since January 1961.

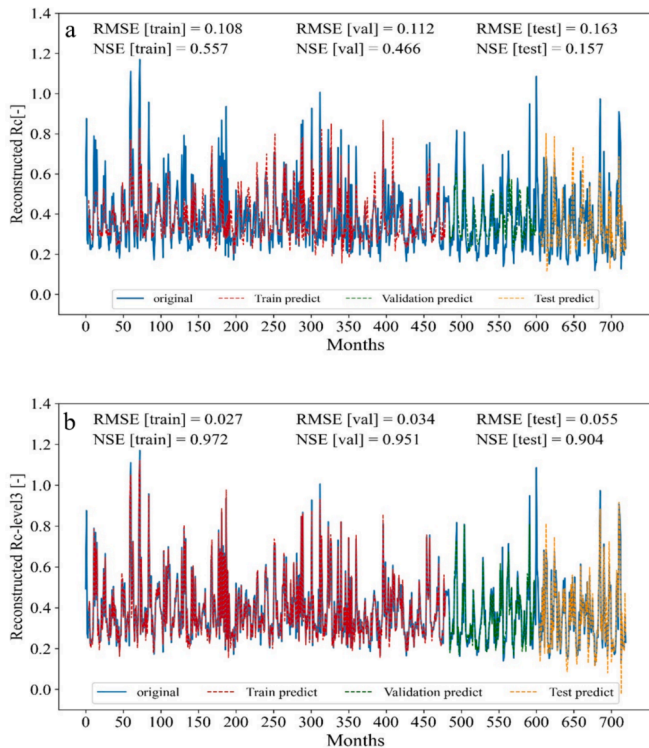


Fig. 7. Results (RMSE and NSE) of fitting a) modelled Rc without decomposition b) a reconstructed Rc modelled at level 3 of decomposition. Months since January 1961.

(Fig. 7a). To overcome this challenge, Rc was decomposed to specific levels of 1 and 3, with particular emphasis on employing the LSTM model through the decomposed components of each level. Then

reconstructed Rc was obtained from equation (7) for both levels 1 and 3 of decomposition. The results show that NSE values of decomposed Rc at level 3 are enhanced significantly in comparison with undecomposed Rc, mainly regarding the test set which shows a value of 0.90 being much higher compared to the value of 0.15 for that of the test dataset of undecomposed signal. Additionally, the results of level 3 of decomposition remain almost constant compared to level 1 except for the test set where it improves. This enhancement may refer to decomposing the noise in two components (D1 and D2) which allows the LSTM to better fit the noise components.

The LSTM model has been successfully implemented across different levels of decomposition, demonstrating very good accuracy. In contrast, the LSTM fails to fit the undecomposed Rc adequately, yielding less acceptable accuracy. This underscores the effectiveness of leveraging decomposition techniques in conjunction with the LSTM model for improved accuracy in modelling the Rc. This indicates the importance of considering the complex details captured at various decomposition levels. However, the findings show that the accuracy of the developed models remained nearly constant for levels beyond 3.

The integration of ERA5 land service data with observed data in the LSTM model showcases remarkable resilience (Pelosi et al., 2020; Muñoz-Sabater et al., 2021), as the model's accuracy and simulation performance remain unaffected. This seamless combination underscores the robustness of the LSTM framework in accommodating diverse data sources, reaffirming its reliability in capturing the intricacies of hydrological dynamics. The methodology presented in this paper can indeed be extended to larger scales, though it is essential to consider several factors. One of the primary advantages of this approach is its reliance on widely available meteorological and hydrological data, which makes it feasible to apply in various regions, even if only small amount of high-resolution data is available. However, when scaling up to regional or global levels, challenges such as data heterogeneity, varying climatic conditions, and differing land-use patterns must be addressed. Moreover, in highly anthropized basins, where human activities significantly alter the hydrological cycle, additional factors such as water

withdrawals, reservoir regulations, and land-use changes must be incorporated into the model. This necessitates more complex modelling frameworks that can account for these variables. Similarly, in regions with limited data availability, the accuracy of Rc predictions may be reduced. In such cases, employing satellite data and integrating multiple data sources can help improve model performance. Overall, while our model provides a robust framework for Rc estimation, careful consideration of local conditions and data quality is essential when applying it to different scales and contexts. Future work should focus on enhancing model adaptability and accuracy in diverse and data-limited environments.

To better visualize the results of modelling Rc at level 3 of decomposition, 3 insets have been illustrated in Fig. B1. These insets (Panels b-d of Fig. B1) refer to the windows among different sets to show more detailed information about the fitted sets. The results indicate a reduction in peak underestimation when combining the LSTM and decomposition method compared to the results of undecomposed Rc.

5. Conclusions

This study introduces an innovative LSTM modelling approach for the prediction of Rc on the basis of key hydrometeorological variables, employing the Rur basin in Germany as a case study for the period 1961 to 2021. The findings reveal that the standard LSTM model performed inadequately when applied to undecomposed Rc, yielding unsatisfactory accuracy across training, validation, and test sets. However, although optimizing the look-back window to 4 months notably improved the modelling in comparison to other values (1–6 months lag), it was still a challenge to predict Rc with acceptable accuracy. Using the LSTM model in conjunction with Maximal overlap Discrete Wavelet Transform demonstrated effectiveness across various levels of decomposition. In our study we decomposed the Rc signal in three levels spanning frequency levels of 0–0.061 and 0.056–0.500 cycle/month for approximation and details, respectively. The NSE values of 0.97, 0.95, and 0.90 for the training, validation, and test sets at level 3 of decomposition, underscore the robust performance of this approach. Notably, the results showcase a very good accuracy of decomposition approach compared to

undecomposed Rc. Additionally, we found an approximately 6 % enhancement in the NSE value of the test set when modelling at level 3 compared to level 1, while the NSE values for the training and validation sets remained constant. These findings suggest the potential application of this innovative technique through hydrological modelling to enhance the reliability of prediction in future studies.

CRedit authorship contribution statement

Arash Rahi: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mehdi Rahmati:** Writing – review & editing, Validation, Supervision, Software, Methodology, Investigation, Conceptualization. **Jacopo Dari:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Investigation, Conceptualization. **Heye Bogena:** Writing – review & editing, Validation, Supervision, Methodology, Data curation. **Harry Vereecken:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Renato Morbidelli:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors wish to thank Leonardo Magrini for the technical support. The authors acknowledge the Climate Data Center (CDC) of the German weather service for providing part of the data used in this study.

Appendix A. WCA between Rc and input variables

Rahi et al. 2023 employed the continuous WCA to assess the distribution of correlation between Rc and hydroclimatic variables. The normalized cross-wavelet spectrum, known as the coherence coefficient being localized in time, highlights the maximum correlation between the two signals (Graf et al., 2014). In our study, we assessed the phase shift and lag between the benchmark signal (Rc) and predictors (i.e., P, SWS, T, and ET). The phase shift between the base signal (Rc) and the second signal (e.g., SWS) is quantified as follows according to Rahmati et al. (2020):

$$\text{Phaseshift} = \frac{\text{Phaseangle}}{2\pi} \times m \quad (\text{A.1})$$

where, m is the period of signals, and the phase shift (in months) and phase angle (in radians) represent the timing difference between consecutive maximal values of the base and second signals. To assess the control relationship between Rc and hydroclimatic parameters, the concept of “lag” defined as follows:

$$\begin{cases} \text{lag} = \text{Phaseshift}, \text{ if } -\frac{m}{4} \leq \text{Phaseshift} \leq \frac{m}{4} \\ \text{lag} = \frac{m}{2} - |\text{Phase shift}|, \text{ if } \text{Phaseshift} < -\frac{m}{4} \\ \text{lag} = \text{Phaseshift} - \frac{m}{2}, \text{ if } \text{Phaseshift} > \frac{m}{4} \end{cases} \quad (\text{A.2})$$

Negative lag values implicitly indicate that the second signal controls the base signal, while positive lag values signify the reverse.

WCA indicates that the correlation between Rc and P is erratic and unstable (see panel (a) of Fig. A1), and in most cycles (see panel (e) of Fig. A1) correlation is insignificant (Coherence < 0.5). This suggests that basin-wide seasonality is not predominant, and Rc is more influenced by water management and human-induced climate change interventions. Scale-averaged phase shift plot shows that P lags 0–5 months behind the Rc, while the negative values of lag indicate that P controls the Rc in the whole time span except for 2011–2015 (panel (c) of Fig. A1). Panel (d) of Fig. A1 depicts increasing lag and phase shift with cycle increments, with the lag showing a lesser increase compared to the phase shift.

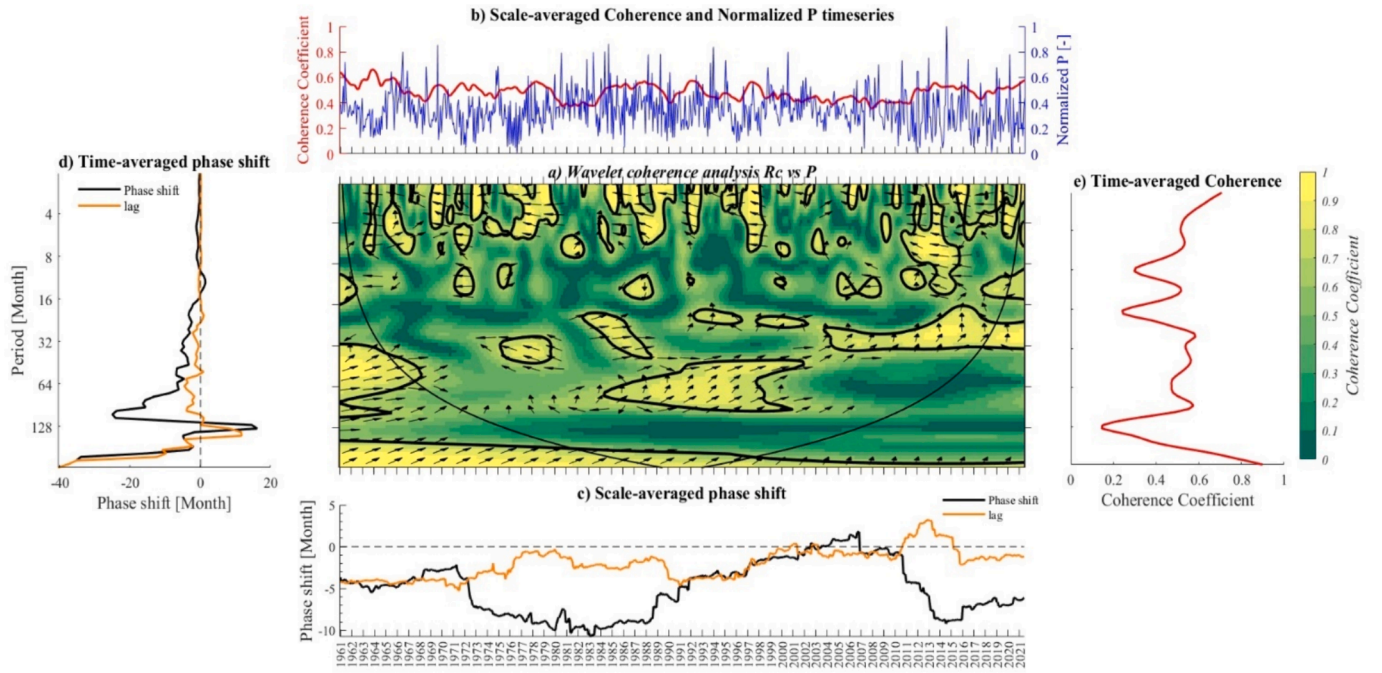


Fig. A1. WCA (a) between Rc and P, (b) scale-averaged coherence and normalized P time series, (c) scale-averaged phase shift, (d) time-averaged phase shift, and (e) time-averaged coherence in Rur basin at Stah outlet from 1961 to 2021.

WCA reveals a notable positive correlation between Rc and SWS in the annual cycle (see panel (a) of Fig. A2). In panel (e) of Fig. A2, as a coefficient of 0.8 is observed. Panel (b) of Fig. A2 captures the 1976–1978 drought, as one of the most-cited drought event (Müller, 2020), resulting in a significant decrease in SWS to around 0.8 m, which notably influences the Rc-SWS correlation. Panel (d) of Fig. A2 indicates minimal lag in cycles shorter than a year, with increasing phase shift and lag values as the cycle lengths. Furthermore, panel (c) of Fig. A2 indicates a lag of 0–5 months, meaning that SWS lags behind the Rc. The lag value unveils a detailed description of evolving dynamics within the system. From 1961 to 2011, negative values indicate that SWS controlled Rc, but in the subsequent period (2011–2021), with positive lag value Rc assumed dominance, showcasing a complex interaction. This dynamic shift adds depth to our understanding of hydrological processes, highlighting adaptability and reciprocal influence.

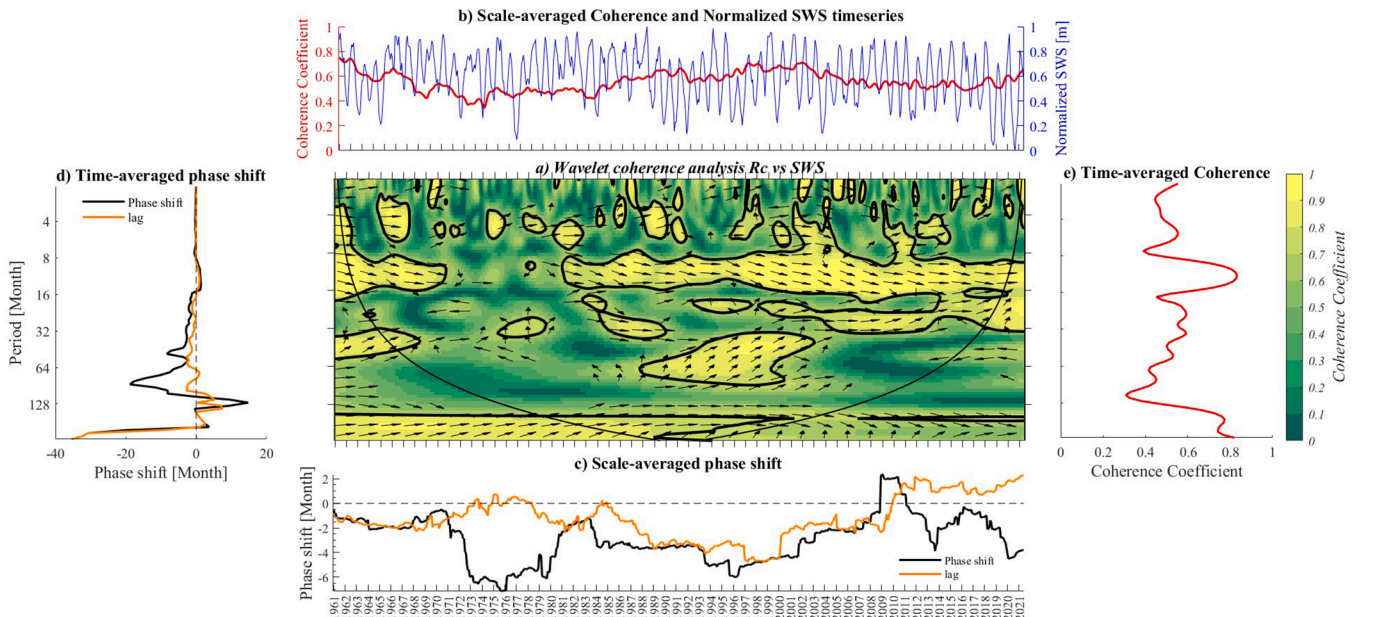


Fig. A2. WCA (a) between Rc and SWS, (b) scale-averaged coherence and normalized SWS time series, (c) scale-averaged phase shift (d) time-averaged phase shift (e) time-averaged coherence in Rur basin at Stah outlet from 1961 to 2021.

WCA shows a noteworthy negative correlation between Rc and ET in the annual cycle (see panel (a) of Fig. A3), and the coherence coefficient attains a substantial value of 0.9 during this cycle (see panel (e) of Fig. A3). The dynamic relationship between Rc and ET is highlighted by varying lag patterns, notably during significant periods from 1980 to 2012. Examining panel (c) of Fig. A3, exhibits that ET lags behind Rc by 0 to 5 months throughout the entire time span, except during 1980–1992 and 2003–2012. Notably, during these intervals with negative lag values, ET exerts control over Rc. Contrarily, panel (d) of Fig. A3 illustrates no lag in the cycle less than annual, indicating synchronised behaviour between Rc and ET considering the time-average frame.

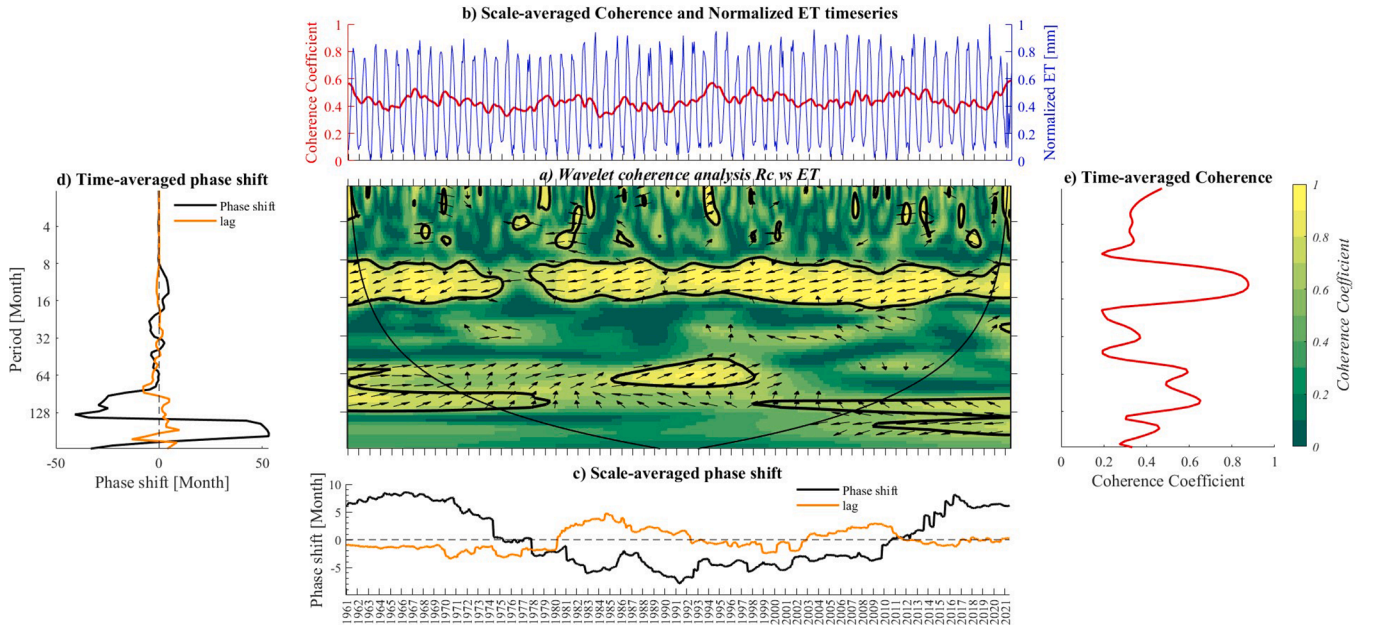


Fig. A3. WCA (a) between R_c and ET, (b) scale-averaged coherence and normalized ET time series, (c) scale-averaged phase shift, (d) time-averaged phase shift, and (e) time-averaged coherence in Rur basin at Stah outlet from 1961 to 2021.

WCA points out a not-negligible negative correlation between the R_c and T within the annual cycle (see panel (a) of Fig. A4). The coherence coefficient attains a considerable value of 0.9 during this annual cycle (see panel (e) of Fig. A4). Further insights from panel (c) of Fig. A4 reveal that T lags 0–6 months behind R_c . Across the vast majority of the observed time span, T consistently lags behind R_c , exercising its influence over the R_c , however, the recent years mark a shift in this established trend. Examination of the time-averaged plot underscores a consistent minor lag present in nearly all cycles (see panel (d) of Fig. A4), shedding light on temperature's proactive role in shaping the dynamics of R_c . This relationship emphasises the complex interplay between T and the R_c , providing valuable insights into the hydrological response to T variations over time.

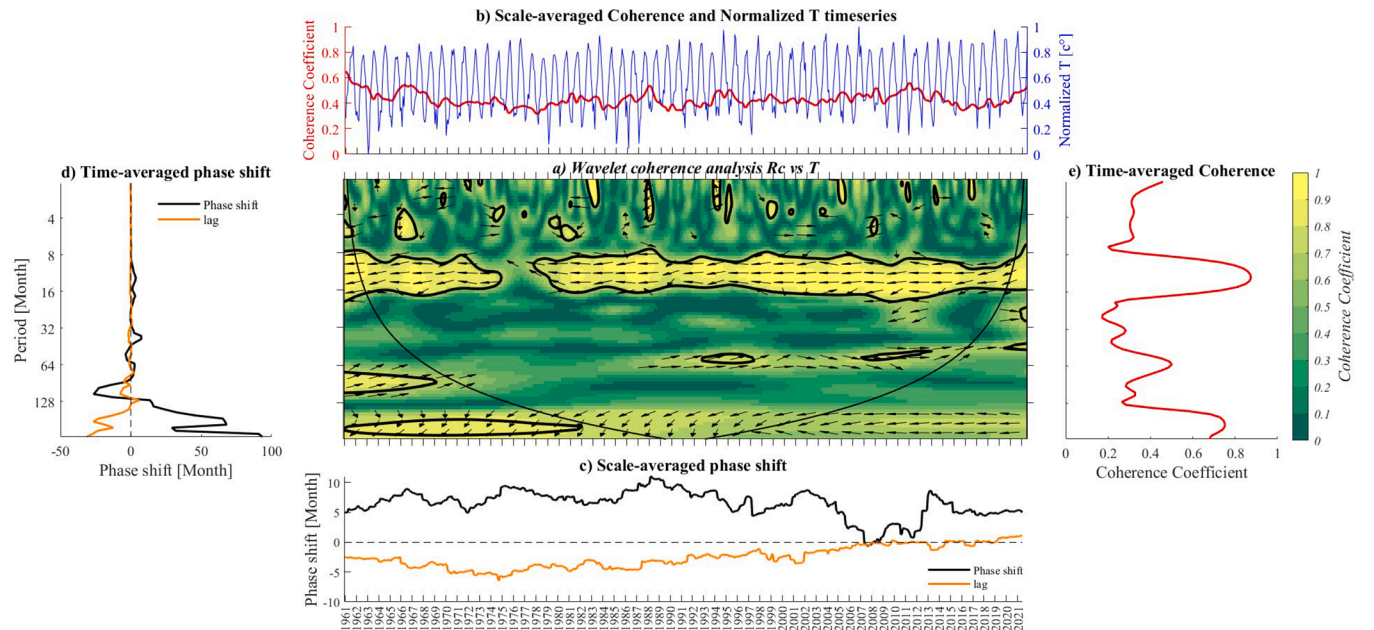


Fig. A4. WCA (a) between R_c and T, (b) scale-averaged coherence and normalized T time series, (c) scale-averaged phase shift, (d) time-averaged phase shift, and (e) time-averaged coherence in Rur basin at Stah outlet from 1961 to 2021.

Appendix B. A closer look to the result of model accuracy at level 3 of decomposition

The results presented in Fig. B1 highlight the improved performance of the model in forecasting R_c at level 3 of decomposition by combining MODWT with LSTM. Panel a shows the overall comparison of the undecomposed R_c values (blue solid line) with the predicted values from the train (red dashed line), validation (green dashed line), and test (yellow dashed line) sets, indicating a general alignment. Panels b, c, and d provide a detailed view of the model's performance on the training, validation, and test sets, respectively. In these insets, the close alignment between the predicted and undecomposed R_c values underscores the model's ability to accurately capture peaks and overall trends. This improvement demonstrates the effectiveness of integrating MODWT with LSTM in enhancing the model's capability to learn and predict complex patterns in the data.

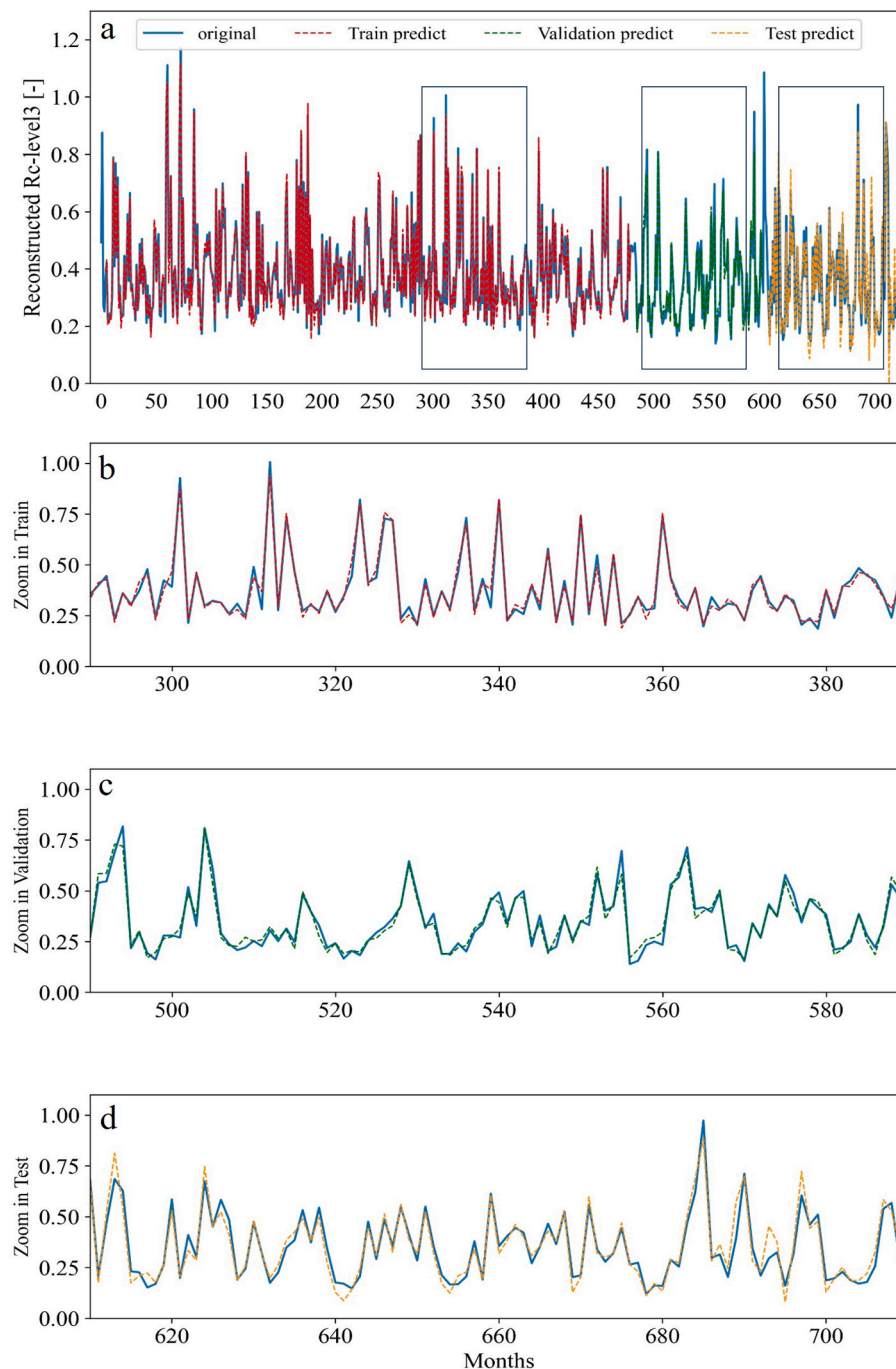


Fig. B1. a) A reconstructed R_c modelled at level 5 of decomposition, insets of the b) train set from March 1985 to August 1993, c) validation set from November 2001 to April 2010, and d) test set from November 2011 to April 2020.

References

- Adamowski, J., Sun, K., 2010. Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. *J. Hydrol.* 390 (1–2) <https://doi.org/10.1016/j.jhydrol.2010.06.033>.
- Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., Harmel, R.D., van Griensven, A., van Liew, M.W., Kannan, N., Jha, M.K., 2012. SWAT: model use, calibration, and validation. *Trans. ASABE* 55 (4).
- Bedient, P.B., Huber, W.C., Vieux, B.E. 2008. *Hydrology and floodplain analysis*. Upper Saddle River, Vol. 816, NJ: Prentice Hall.
- Bergen, K.J., Johnson, P.A., de Hoop, M.v., Beroza, G.C., 2019. Machine learning for data-driven discovery in solid Earth geoscience. *Science* 363 (6433). <https://doi.org/10.1126/science.aau0323>.
- Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A.L., Deng, D., Lindauer, M., 2023. Hyperparameter optimization: foundations, algorithms, best practices, and open challenges. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 13 (2) <https://doi.org/10.1002/widm.1484>.
- Blöschl, G., Hall, J., Viglione, A., Perdigão, R.A.P., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G.T., Bilibashi, A., Boháč, M., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G.B., Claps, P., Frolova, N., Ganora, D., Živković, N., 2019. Changing climate both increases and decreases European river floods. *Nature* 573 (7772). <https://doi.org/10.1038/s41586-019-1495-6>.
- Bogena, H., Hake, J.F., Herbst, M., Kunkel, R., Montzka, C., Pütz, T., et al., 2005a. MOSYRUR: Water balance analysis in the Rur basin. *Schr. Forschungszent. Juelich, Reihe Umwelt/Environ.* 52. Forschungszentrum Jülich, Jülich, Germany.
- Bogena, H., Kunkel, R., Schöbel, T., Schrey, H.P., Wendland, F., 2005b. Distributed modeling of groundwater recharge at the macroscale. *Ecol. Model.* 187 (1 SPEC. ISS.) <https://doi.org/10.1016/j.ecolmodel.2005.01.023>.

- Burak, S., Bilge, A.H., Ulker, D., 2021. Computation of monthly runoff coefficients For Istanbul. *Therm. Sci.* 25 (2 Part B), 1561–1572. <https://doi.org/10.2298/TSC191102147B>.
- Chen, J., Kumar, P., 2002. Role of terrestrial hydrologic memory in modulating ENSO impacts in North America. *J. Clim.* 15 (24) [https://doi.org/10.1175/1520-0442\(2003\)015<3569:ROTHMI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)015<3569:ROTHMI>2.0.CO;2).
- Chen, X., Parajka, J., Széles, B., Strauss, P., Blöschl, G., 2020. Spatial and temporal variability of event runoff characteristics in a small agricultural catchment. *Hydrol. Sci. J.* <https://doi.org/10.1080/02626667.2020.1798451>.
- Chen, Y., Xu, J., 2021. Rainfall-runoff short-term forecasting method based on LSTM. *J. Phys. Conf. Ser.* 2025 (1) <https://doi.org/10.1088/1742-6596/2025/1/012005>.
- Chollet, F. (2021). Keras: The Python Deep Learning library. <https://keras.io>.
- Cornish, C.R., Bretherton, C.S., Percival, D.B., 2006. Maximal overlap wavelet statistical analysis with application to atmospheric turbulence. *Boundary-Layer Meteorol.* 119 (2) <https://doi.org/10.1007/s10546-005-9011-y>.
- Dari, J., Flammini, A., Morbidelli, R., Rahi, A., Saltalippi, C., 2023. Evolution of freshwater availability in a climate-changing Mediterranean context: the case of Umbria region, central Italy. *Hydrol. Processes J.* 37 (12) <https://doi.org/10.1002/hyp.15050>.
- Dari, J., Quintana-Seguí, P., Barella-Ortiz, A., Rahmati, M., Saltalippi, C., Flammini, A., Brocca, L., 2024. Quantifying the hydrological impacts of irrigation on a Mediterranean agricultural context through explicit satellite-derived irrigation estimates. *Water Resour. Res.* 60 (5) <https://doi.org/10.1029/2023WR036510>.
- Dragomireskiy, K., Zosso, D., 2014. Variational mode decomposition. *IEEE Trans. Signal Process.* 62 (3) <https://doi.org/10.1109/TSP.2013.2288675>.
- Fan, H., Jiang, M., Xu, L., Zhu, H., Cheng, J., Jiang, J., 2020. Comparison of long short term memory networks and the hydrological model in runoff simulation. *Water (Switzerland)* 12 (1). <https://doi.org/10.3390/w12010175>.
- Gelfan, A.N., Semenov, V.A., Motovilov, Y.G., 2015. Climate noise effect on uncertainty of hydrological extremes: numerical experiments with hydrological and climate models. *IAHS-AISH Proc. Rep.* 369 <https://doi.org/10.5194/piahs-369-49-2015>.
- Gharari, S., Razavi, S., 2018. A review and synthesis of hysteresis in hydrology and hydrological modeling: memory, path-dependency, or missing physics? In: *J. Hydrol.* 566 <https://doi.org/10.1016/j.jhydrol.2018.06.037>.
- Girihagama, L., Naveed Khaliq, M., Lamontagne, P., Perdikaris, J., Roy, R., Sushama, L., Elshorbagy, A., 2022. Streamflow modelling and forecasting for Canadian watersheds using LSTM networks with attention mechanism. *Neural Comput. Appl.* 34 (22) <https://doi.org/10.1007/s00521-022-07523-8>.
- Goel, V., Jangir, V., Shankar, V.G., 2020. DataCan: robust approach for genome cancer data analysis. *Adv. Intell. Signal Comput.* 1016 https://doi.org/10.1007/978-981-13-9364-8_12.
- Gokhale, M.Y., Khanduja, D.K., 2010. Time domain signal analysis using wavelet packet decomposition approach. *Int. J. Commun. Netw. Syst. Sci.* 3 (3) <https://doi.org/10.4236/ijcns.2010.33041>.
- Graf, A., Bogen, A., Rüde, C., Hardelauf, H., Pütz, T., Heinemann, G., Vereecken, H., 2014. Spatiotemporal relations between water budget components and soil water content in a forested tributary catchment. *Water Resour. Res.* 50 (6), 4837–4857. <https://doi.org/10.1002/2013WR014516>.
- Guo, Z., Zhang, Q.Q., Li, N., Zhai, Y.Q., Teng, W.T., Liu, S.S., Ying, G.G., 2023. Runoff time series prediction based on hybrid models of two-stage signal decomposition methods and LSTM for the Pearl River in China. *Hydrol. Res.* 54 (12) <https://doi.org/10.2166/nh.2023.069>.
- Gusev, Y.M., Semenov, V.A., Nasonova, O.N., Kovalev, E.E., 2017. Weather noise impact on the uncertainty of simulated water balance components of river basins. *Hydrol. Sci. J.* 62 (8) <https://doi.org/10.1080/02626667.2017.1319064>.
- Hassani, V., Tjahjowidodo, T., Do, T.N., 2014. A survey on hysteresis modeling, identification and control. *Mech. Syst. Signal Process.* 49 (1–2) <https://doi.org/10.1016/j.ymssp.2014.04.012>.
- Hochreiter, S., Schmidhuber, J., 1997. Long short term memory. *Neural Comput.* 9 (8), 1735–1780.
- Holman-Dodds, J.K., Bradley, A.A., Sturdevant-Rees, P.L., 1999. Effect of temporal sampling of precipitation on hydrologic model calibration. *J. Geophys. Res. Atmos.* 104 (D16) <https://doi.org/10.1029/1999JD900121>.
- Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Yen, N., Tung, C.C., Liu, H.H., 1996. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. Royal Soc. A* 454 (1971).
- Kingma, D.P., Ba, J.L., 2015. Adam: A method for stochastic optimization. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.
- Kratzert, F., Klotz, D., Brenner, C., Schulz, K., Herrnegger, M., 2018. Rainfall-runoff modelling using long short-term memory (LSTM) networks. *Hydrol. Earth Syst. Sci.* 22 (11) <https://doi.org/10.5194/hess-22-6005-2018>.
- Li, J., Yang, B., Li, H., Wang, Y., Qi, C., Liu, Y., 2021. DTDR-ALSTM: Extracting dynamic time-delays to reconstruct multivariate data for improving attention-based LSTM industrial time series prediction models. *Knowl.-Based Syst.* 211, 106508 <https://doi.org/10.1016/j.knsys.2020.106508>.
- Liu, Z., Zhou, P., Chen, G., Guo, L., 2014. Evaluating a coupled discrete wavelet transform and support vector regression for daily and monthly streamflow forecasting. *J. Hydrol.* 519 (PD) <https://doi.org/10.1016/j.jhydrol.2014.06.050>.
- Mallat, S.G., 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (7), 674–693.
- Mao, G., Wang, M., Liu, J., Wang, Z., Wang, K., Meng, Y., Zhong, R., Wang, H., Li, Y., 2021. Comprehensive comparison of artificial neural networks and long short-term memory networks for rainfall-runoff simulation. *Phys. Chem. Earth* 123.
- Massari, C., Pellet, V., Trambly, Y., Crow, W.T., Gründemann, G.J., Hascote, T., Penna, D., Modanesi, S., Brocca, L., Camici, S., Marra, F., 2023. On the relation between antecedent basin conditions and runoff coefficient for European floods. *J. Hydrol.* 625 <https://doi.org/10.1016/j.jhydrol.2023.130012>.
- Masseroni, D., Camici, S., Cislighi, A., Vacchiano, G., Massari, C., Brocca, L., 2021. The 63-year changes in annual streamflow volumes across Europe with a focus on the Mediterranean basin. *Hydrol. Earth Syst. Sci.* 25 (10) <https://doi.org/10.5194/hess-25-5589-2021>.
- Merz, R., Blöschl, G., Parajka, J., 2006. Spatio-temporal variability of event runoff coefficients. *J. Hydrol.* 331 (3–4), 591–604. <https://doi.org/10.1016/j.jhydrol.2006.06.008>.
- Merz, R., Blöschl, G., 2009. A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. *Water Resour. Res.* 45 (1) <https://doi.org/10.1029/2008WR007163>.
- Moosavi, V., Talebi, A., Hadian, M.R., 2017. Development of a hybrid wavelet packet-group method of data handling (WPGMDH) model for runoff forecasting. *Water Resour. Manag.* 31 (1) <https://doi.org/10.1007/s11269-016-1507-3>.
- Moriasi, D.N., Gitau, M.W., Pai, N., Dagguapati, P., 2015. Hydrologic and water quality models: performance measures and evaluation criteria. *Trans. ASABE* 58 (6). <https://doi.org/10.13031/trans.58.10715>.
- Müller, W., 2020. Drought victims demand justice: politicization of drought by farmers in Southern Germany over time. *Water (Switzerland)* 12 (3). <https://doi.org/10.3390/w12030871>.
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* 13 (9). <https://doi.org/10.5194/essd-13-4349-2021>.
- Naabil, E., Lamptey, B.L., Arnault, J., Kunstmann, H., Olufayo, A., 2017. Water resources management using the WRF-Hydro modelling system: case-study of the Tono dam in West Africa. *J. Hydrol.: Reg. Stud.* 12 <https://doi.org/10.1016/j.ejrh.2017.05.010>.
- Napolitano, G., Serinaldi, F., See, L., 2011. Impact of EMD decomposition and random initialization of weights in ANN hindcasting of daily stream flow series: an empirical examination. *J. Hydrol.* 406 (3–4) <https://doi.org/10.1016/j.jhydrol.2011.06.015>.
- Nhita, F., Saepudin, D., Adiwijaya, Wisesty, U.N., 2016. Comparative Study of Moving Average on Rainfall Time Series Data for Rainfall Forecasting Based on Evolving Neural Network Classifier. *Proceedings - 2015 3rd International Symposium on Computational and Business Intelligence, ISCBI 2015*. <https://doi.org/10.1109/ISCBI.2015.27>.
- Norbiato, D., Borgia, M., Merz, R., Blöschl, G., Carton, A., 2009. Controls on event runoff coefficients in the eastern Italian Alps. *J. Hydrol.* 375 (3–4), 312–325. <https://doi.org/10.1016/j.jhydrol.2009.06.044>.
- Nourani, V., Kisi, Ö., Komasi, M., 2011. Two hybrid artificial intelligence approaches for modeling rainfall-runoff process. *J. Hydrol.* 402 (1–2) <https://doi.org/10.1016/j.jhydrol.2011.03.002>.
- Onstad, C.A., Jamieson, D.G., 1970. Modeling the effect of land use modifications on runoff. *Water Resour. Res.* 6 (5) <https://doi.org/10.1029/WR006i005p01287>.
- Orth, R., Seneviratne, S.I., 2013. Propagation of soil moisture memory to streamflow and evapotranspiration in Europe. *Hydrol. Earth Syst. Sci.* 17 (10) <https://doi.org/10.5194/hess-17-3895-2013>.
- Patry, G.G., Kennedy, A., Potter, S., 1989. Runoff modelling under noise-corrupted rainfall conditions. *Can. J. Civ. Eng.* 16 (5) <https://doi.org/10.1139/89-100>.
- Pelosi, A., Terribile, F., D'Urso, G., Chirico, G.B., 2020. Comparison of ERA5-Land and UERRA MESCAN-SURFEX reanalysis data with spatially interpolated weather observations for the regional assessment of reference evapotranspiration. *Water (Switzerland)* 12 (6). <https://doi.org/10.3390/W12061669>.
- Penna, D., Tromp-van Meerveld, H.J., Gobbi, A., Borgia, M., Dalla Fontana, G., 2011. The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment. *Hydrol. Earth Syst. Sci.* 15 (3), 689–702. <https://doi.org/10.5194/hess-15-689-2011>.
- Piotrowski, A.P., Napiorkowski, J.J., 2013. A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling. *J. Hydrol.* 476 <https://doi.org/10.1016/j.jhydrol.2012.10.019>.
- Rahi, A., Rahmati, M., Dari, J., Saltalippi, C., Brogi, C., Morbidelli, R., 2023. Unraveling hydroclimatic forces controlling the runoff coefficient trends in central Italy's Upper Tiber Basin. *J. Hydrol.: Reg. Stud.* 50, 2214–5818. <https://doi.org/10.1016/j.ejrh.2023.101579>.
- Rahmati, M., Groh, J., Graf, A., Pütz, T., Jan Vanderborcht, J., Vereecken, H., 2020. On the impact of increasing drought on the relationship between soil water content and evapotranspiration of a grassland. *Vadose Zone J.* 19, e20029.
- Ranjan, G.S.K., Kumar Verma, A., Radhika, S., 2019. K-Nearest Neighbors and Grid Search CV Based Real Time Fault Monitoring System for Industries. In: 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), pp. 1–5. <https://doi.org/10.1109/I2CT45611.2019.9033691>.
- Rudi, J., Pabel, R., Jager, G., Koch, R., Kunoth, A., Bogen, H., 2010. Multiscale analysis of hydrologic time series data using the Hilbert-Huang transform. *Vadose Zone J.* 9 (4) <https://doi.org/10.2136/vzj2009.0163>.
- Sáez, J.A., Galar, M., Luengo, J., Herrera, F., 2014. Analyzing the presence of noise in multi-class problems: alleviating its influence with the one-vs-one decomposition. *Knowl. Inf. Syst.* 38 (1) <https://doi.org/10.1007/s10115-012-0570-1>.
- Seo, Y., Choi, Y., Choi, J., 2017. River stage modeling by combining maximal overlap discrete wavelet transform, support vector machines and genetic algorithm. *Water (Switzerland)* 9 (7). <https://doi.org/10.3390/w9070525>.
- Sherman, L.K., 1932. On runoff. *Trans. Am. Geophys. Union* 13 (1), 298. <https://doi.org/10.1029/TR013i001p00298-1>.
- Shoaib, M., Shamseldin, A.Y., Melville, B.W., 2014. Comparative study of different wavelet based neural network models for rainfall-runoff modeling. *J. Hydrol.* 515 <https://doi.org/10.1016/j.jhydrol.2014.04.055>.

- Shukla, S., Meshesha, T.W., Sen, I.S., Bol, R., Bogen, H., Wang, J., 2023. Assessing impacts of land use and land cover (LULC) change on stream flow and runoff in Rur Basin, Germany. *Sustainability (Switzerland)* 15 (12). <https://doi.org/10.3390/su15129811>.
- Srivastava, M., Anderson, C.L., Freed, J.H., 2016. A new wavelet denoising method for selecting decomposition levels and noise thresholds. *IEEE Access* 4, 3862–3877. <https://doi.org/10.1109/ACCESS.2016.2587581>.
- Velpuri, N.M., Senay, G.B., 2013. Analysis of long-term trends (1950–2009) in precipitation, runoff and runoff coefficient in major urban watersheds in the United States. *Environ. Res. Lett.* 8 (2), 024020.
- Wang, W.C., Chau, K.W., Qiu, L., Chen, Y.B., 2015. Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. *Environ. Res.* 139 <https://doi.org/10.1016/j.envres.2015.02.002>.
- Wang, Y., Gao, J., Xu, Z., Li, L., 2020. A short-term output power prediction model of wind power based on deep learning of grouped time series. *Eur. J. Electr. Eng.* 22 (1) <https://doi.org/10.18280/ejee.220104>.
- Wu, C.L., Chau, K.W., Li, Y.S., 2009. Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques. *Water Resour. Res.* 45 (8) <https://doi.org/10.1029/2007WR006737>.
- Xiang, Z., Yan, J., Demir, I., 2020. A rainfall-runoff model with LSTM-based sequence-to-sequence learning. *Water Resour. Res.* 56 (1).
- Zacharias, S., Loescher, H., Bogen, H., Kiese, R., Schrön, M., Vereecken, H., 2024. 15 years of integrated terrestrial environmental observatories (TERENO) in Germany: functions, services and lessons learned. *ESS Open Archive*. <https://doi.org/10.22541/essoar.170808463.36288013/v1>.
- Zhou, H., Schertzer, D., Tchiguirinskaia, I., 2024. Combining recurrent neural networks with variational mode decomposition and multifractals to predict rainfall time series. *EGU sphere* [preprint]. <https://doi.org/10.5194/egusphere-2023-2710>.